

University of Nebraska - Lincoln

**DigitalCommons@University of Nebraska - Lincoln**

---

Publications, Agencies and Staff of the U.S.  
Department of Commerce

U.S. Department of Commerce

---

2002

## Application and National Geographic Information System Database to Support Two-Year Flood and Threshold Runoff Estimates

Seann Reed

*National Weather Service Hydrology Laboratory*

Dennis Johnson

*Juniata College*

Timothy Sweeney

*National Weather Service Hydrology Laboratory*

Follow this and additional works at: <https://digitalcommons.unl.edu/usdeptcommercepub>

 Part of the [Environmental Sciences Commons](#)

---

Reed, Seann; Johnson, Dennis; and Sweeney, Timothy, "Application and National Geographic Information System Database to Support Two-Year Flood and Threshold Runoff Estimates" (2002). *Publications, Agencies and Staff of the U.S. Department of Commerce*. 55.  
<https://digitalcommons.unl.edu/usdeptcommercepub/55>

This Article is brought to you for free and open access by the U.S. Department of Commerce at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Publications, Agencies and Staff of the U.S. Department of Commerce by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Application and National Geographic Information System Database to Support Two-Year Flood and Threshold Runoff Estimates

Seann Reed<sup>1</sup>; Dennis Johnson<sup>2</sup>; and Timothy Sweeney<sup>3</sup>

**Abstract:** A computer application and national geospatial database have been developed to support the calculation of *flooding flow* ( $Q_f$ ) and threshold runoff across the conterminous United States and Alaska. Flooding flow is the flow required to cause a stream to slightly overflow its bank and cause damage. Threshold runoff [ $L$ ], defined as the depth of runoff required to cause flooding, is computed as flooding flow divided by the unit hydrograph peak flow. A key assumption in this work is that the two-year return flood ( $Q_2$ ) is a useful surrogate for flooding flow. The application described here computes flood magnitude estimates for selected return periods ( $Q_2$ ,  $Q_5$ ,  $Q_{10}$ , etc.) using regression equations published by the U.S. Geological Survey for each of 210 hydrologic regions. The application delineates basin boundaries and computes all basin parameters required for the flood frequency calculations. The geographic information system database that supports these calculations contains terrain data [digital elevation models (DEMs) and DEM derivatives], reference data, and 89 additional data layers related to climate, soils, geology, and land use. Initial results indicate that there are some practical limitations associated with using  $Q_2$  regression equations to estimate flooding flow.

**DOI:** 10.1061/(ASCE)1084-0699(2002)7:3(209)

**CE Database keywords:** Geographic information systems; Alaska; Databases; and Runoff.

## Introduction

Flooding is a natural and necessary part of the ecosystem of a river. A flood occurs when the accumulation of water exceeds the carrying capacity of the channel system. This is generally due to either the rainfall rate exceeding the potential infiltration rate or the cumulative infiltration exceeding the storage capacity of the watershed. Thus, both the intensity and the duration of precipitation are key factors in determining the severity of a flood event. One of the characteristic and distinguishable elements of a flood is the timing or time-to-peak.

A flash flood can be defined as a flood that peaks within 6 h of heavy rainfall (Sweeney 1992). Flash floods often occur in situations when the rainfall intensity exceeds the soil infiltration rate, and typically occur in basins smaller than 259 km<sup>2</sup> (100 mi<sup>2</sup>) (Davis 1998). There may be many reasons for rainfall intensity to exceed the maximum soil infiltration rate, including high antecedent rainfall, low permeability soils, and human alteration of the watershed (deforestation, urbanization, etc.), to name a few.

As part of its mission to issue weather, hydrologic, and climate forecasts and warnings, the United States National Weather Service (NWS) has developed a flash flood guidance (*FFG*) system to aid forecasters in assessing when to issue flash flood watches and warnings. The quantity, *FFG*, is defined as the average rainfall depth falling over a given area for a specified duration that is required to cause flooding on small streams (Sweeney 1992). The key aspects of the *FFG* definition are a depth and duration of rainfall for a given location and time. The National Weather Service River Forecast Centers (RFCs) currently compute 1, 3, and 6 h *FFG* at least once per day.

*FFG* for a basin is computed at any given time using the following three pieces of information: (1) a rainfall-runoff curve for a given duration of rainfall that reflects the current basin soil moisture conditions; (2) a basin routing scheme used to translate runoff depth to flow at the basin outlet (unit hydrograph is currently used); and (3) a threshold flow level or flooding flow ( $Q_f$ ) at which overbank or damaging flow occurs. Since, by definition, a basin unit hydrograph for a given duration is time invariant, it is convenient to combine the unit hydrograph [(2)] and flooding flow [(3)] information into a single quantity referred to as threshold runoff. Threshold runoff ( $R$ ) is the depth of runoff [ $L$ ] over a basin for a given duration required to exceed flooding flow at the outlet

$$R = \frac{Q_f}{Q_{pR}} \quad (1)$$

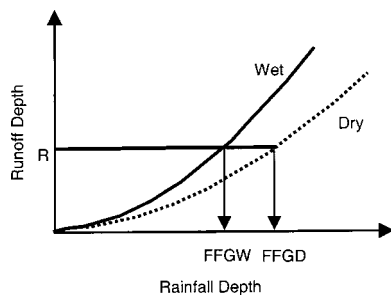
where  $Q_f$ =flooding flow [ $L^3T^{-1}$ ]; and  $Q_{pR}$ =unit hydrograph peak flow [ $L^2T^{-1}$ ] for a given duration (subscript  $R$ ). Different  $Q_{pR}$  values and hence different threshold runoff values are computed for storms of 1, 3, and 6 h duration. Fig. 1 shows how threshold runoff for an area is used in conjunction with a rainfall-runoff curve for that area at a specific time to produce *FFG*. The threshold runoff value for a given location and duration does not

<sup>1</sup>Hydrologist, National Weather Service Hydrology Laboratory, w/OHD1, 1325 East-West Hwy., Silver Spring, MD 20910.

<sup>2</sup>Assistant Professor of Environmental Science and Studies, Juniata College, 1700 Moore St., Huntingdon, PA 16652.

<sup>3</sup>Hydrologist, National Weather Service Hydrology Laboratory, w/OHD1, 1325 East-West Hwy., Silver Spring, MD 20910.

Note. Discussion open until October 1, 2002. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on February 1, 2001; approved on September 24, 2001. This paper is part of the *Journal of Hydrologic Engineering*, Vol. 7, No. 3, May 1, 2002. ©ASCE, ISSN 1084-0699/2002/3-209-219/\$8.00+\$5.50 per page.



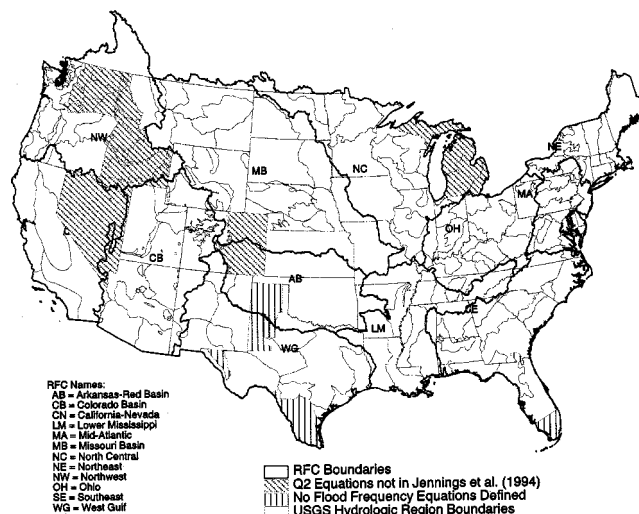
**Fig. 1.** Using threshold runoff ( $R$ ) and current rainfall-runoff curve (wet or dry) to determine flash flood guidance ( $FFG$ )

vary with time, but the rainfall-runoff curve does. The hypothetical curve for “wet” soil conditions in Fig. 1 results in a lower  $FFG$  estimate ( $FFGW$ ) than the curve for “dry” soil conditions ( $FFGD$ ).

Sweeney (1992) identified the need to provide a more consistent and objective tool to RFCs for threshold runoff estimation. To address this requirement, Carpenter and Georgakakos (1993) developed a geographic resources analysis support system (GRASS) based (GRASS 1983) *threshR* software package. Carpenter et al. (1999) describe the use of this software to compute threshold runoff values for most of Iowa and Oklahoma, and part of California. Carpenter et al. (1999) considered both a statistical quantity, the two-year flood ( $Q_2$ ), and a physical quantity, bank-full flow estimated using Manning’s equation, as estimators for flooding flow. Carpenter et al. (1999) also note that bank-full flow is a conservative estimate of flooding flow because flow exceeding bank full is required to cause damage in most locations. Because regional cross-sectional relationships required to estimate bank-full flow using Manning’s equation are not readily available for much of the United States, and equations to estimate  $Q_2$  are available from the U.S. Geological Survey (USGS) (Jennings et al. 1994) for most of the United States, only the use of  $Q_2$  equations is pursued further in this study. The reader is referred to Wolman and Leopold (1957) and Henderson (1966) for further discussion of the return period most closely associated with bank-full flow.

Starting in the spring of 1999, the ArcView-based threshold runoff application described here (*AvThreshR*) was developed to facilitate easier threshold runoff calculations at RFCs. The most noteworthy enhancement in *AvThreshR* is a national database to support  $Q_2$  estimates. The *AvThreshR* database and accompanying programs support automatic calculation of flood magnitudes for different return periods ( $Q_2$ ,  $Q_5$ ,  $Q_{10}$ , etc.) across the conterminous United States and Alaska. The *AvThreshR* programs can efficiently compute required basin characteristics and produce the corresponding  $Q_2$  estimates for thousands of basins at a time. This differs from the National Flood Frequency (NFF) software distributed with Jennings et al. (1994) because the NFF software requires manual entry of basin parameters.

In addition to database and functional support for  $Q_2$  parameter estimation, *AvThreshR* development also included efforts to improve computational efficiency. Digital elevation model (DEM) processing is the most computationally expensive aspect of threshold runoff procedures. Attempts to compute threshold runoff using 3 arcsec DEMs placed a heavy burden on RFC computer systems in the early to mid 1990s. As a result, the following two steps were taken to improve computational efficiency: (1) preprocessing of DEM data sets prior to RFC delivery; and (2) experimentation with coarser, 15 arcsec DEM data. As expected, using the 15 arcsec data in lieu of 3 arcsec data yielded the most



**Fig. 2.** Map of USGS regions used for regression analysis in conterminous United States. Hash marks denote areas where  $Q_2$  rural regression equations are not available from Jennings et al. (1994).

significant speed improvement because 25 times fewer data values are processed, but DEM preprocessing also helped.

There are limitations in using the relatively coarse resolution 15 arcsec data to define parameters for small basins. As a rule of thumb, basins smaller than about  $78 \text{ km}^2$  ( $30 \text{ mi}^2$ ) cannot be accurately delineated using these data (“Basin” 2001). Because Davis (1998) indicates that computing mean areal precipitation on basins as small as  $5 \text{ km}^2$  ( $\sim 2 \text{ mi}^2$ ) can be useful to forecasters, some scale issues associated with threshold runoff calculations are explored in the present paper using a 30-m DEM. Although not addressed in this paper, it is important to recognize that scale issues are also critical to the rainfall-runoff calculations that go into flash flood guidance (Finnerty et al. 1997). Ongoing research into distributed models at the NWS attempts to address this issue (Smith et al. 1999; Koren et al. 2001).

The next two sections of this paper describe the *AvThreshR* database, the computational methods, and the method for a simple uncertainty analysis. This is followed by a discussion of example results and conclusions.

## Data

Jennings et al. (1994) summarize regression equations that can be used to estimate peak floods associated with various recurrence intervals in all 50 U.S. states. Each state summary is derived from one or more earlier publications, which contain the details of the equation derivations. Many states are divided into multiple hydrologic regions where different regression equations apply. There are a total of 210 different hydrologic regions defined in the United States.

Typically, equations for 2 ( $Q_2$ ), 5 ( $Q_5$ ), 10 ( $Q_{10}$ ), 25 ( $Q_{25}$ ), 50 ( $Q_{50}$ ), and 100-year ( $Q_{100}$ ) floods are available for rural areas; however, equations for some of these return periods are not available in all regions. For example, in Michigan, the five-year equation is the shortest duration equation provided in Jennings et al. (1994). Fig. 2 shows areas in the conterminous United States where the  $Q_2$  equations are not available from Jennings et al. (1994). There are also some areas shown in Fig. 2 where the data available to the USGS were inadequate to develop flood-

**Table 1.** Data Sets Common to All River Forecast Centers

Data description	File type	Source
Digital elevation model (m)	Grid	NOHRSC
Flow direction	Grid	Derived
Flow accumulation	Grid	Derived
Downstream flow length (m)	Grid	Derived
DEM cell slope (percent)	Grid	Derived
Buffered mask of RFC boundary (ones inside and NODATA outside)	Grid	Derived
Center points of HRAP cells	Shapefile-point	Created
State boundaries	Shapefile-polygon	USGS
USGS hydrologic region boundaries	Shapefile-polygon	USGS
Modified RF1 file	Shapefile-line	NOHRSC
Parameter information table	Dbase	Created
Regression equation coefficients table	Dbase	Created

frequency equations. The equations for computing  $Q_2$  are of primary interest for threshold runoff calculations; however, computation of  $Q_5$ ,  $Q_{10}$ ,  $Q_{25}$ ,  $Q_{50}$ , or  $Q_{100}$  is also supported by the database where the appropriate regression equations are defined. In some areas of the western United States, peak floods for longer return periods (e.g.,  $Q_5$ ) may provide a more realistic estimate of the flooding flow than  $Q_2$ .

It is important to keep in mind the limitations noted by Jennings et al. (1994) in applying these equations. The equations supported by the *AvThreshR* database do not apply in urban areas, and the presence of dams or other structures that impact peak discharge makes the equations inapplicable. Initial applications of *AvThreshR* are intended to derive representative threshold runoff values for basins without significant anthropogenic influence. Therefore, it is not critical to weed out specific basins with dams; however, a user could easily flag subbasins containing dams using basic ArcView functions without further software development.

The majority of the regression equations take the form shown in

$$Q_T = aX^bY^c \dots \quad (2)$$

where  $a$ ,  $b$ , and  $c$ =constants, and  $X$  and  $Y$ =independent parameters. Up to 12 independent parameters are used in some regions (e.g., Michigan). *AvThreshR* automatically computes the required independent parameters from geographic information system (GIS) data layers. Ten exceptions are incorporated into one of the *AvThreshR* programs to account for the hydrologic regions where equations take a form different from that shown in Eq. (2).

For computational purposes, it is useful to divide the independent parameters into two categories. One category includes parameters that are computed at all locations, and the second category includes parameters that are only computed in specific states. For example, a parameter related to topography like main channel slope is determined using the DEM and its derivatives, and thus can be computed at any location. However, a parameter like soil permeability that is used to estimate  $Q_2$  in Kansas is not required in other states, and therefore is only computed for basins in Kansas. In fact, the soil permeability data layer required for Kansas only covers the state of Kansas.

The *AvThreshR* application is designed for use at each of the 13 RFCs; hence, many of the data layers are organized according to RFC service areas. The RFC service areas are outlined in Fig. 2. Table 1 lists data sets common to all RFCs, while Table 2 lists 14 additional types of data that are used in some, but not all states. For example, 18 of the states require information about the

**Table 2.** State-Specific Data Layers Required to Support Regression Equation Calculations

Data layer description	Number of data layers of this type
Surface water storage (lakes, ponds, swamps)	18
Mean annual precipitation	16
Minimum mean January temperature	4
Rainfall amount for given duration	20
Forest cover	9
Soil characteristics	6
Mean annual snowfall	2
Geological characteristics	8
Runoff coefficient	1
Mean annual runoff	1
Normal daily May–March temperature	1
Impervious cover	1
Annual PET	1
Geographic factor	1

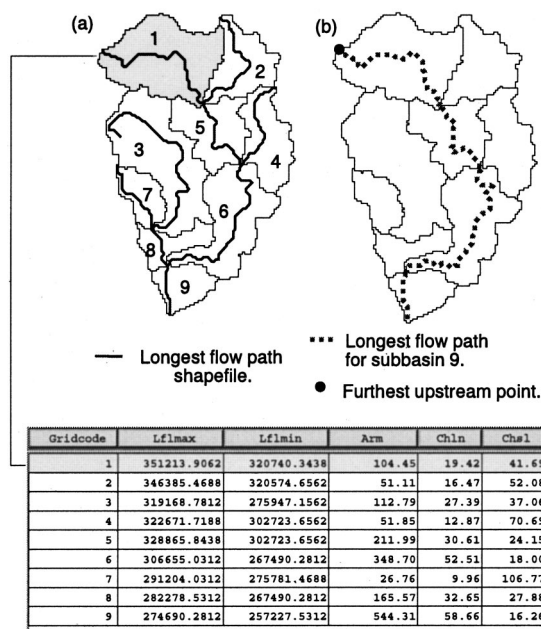
percent of a basin's surface area occupied by water, 16 states require information about mean annual precipitation, and so forth. A total of 89 state-specific layers (with data covering only the state of interest) are included in the database as of October 2000. Data sets listed in Table 1 were obtained from the U.S. Geological Survey, the National Weather Service National Operational Hydrologic Remote Sensing Center (NOHRSC), derived from NOHRSC data sets, or created from scratch. Endres (1999) describes the sources for data layers listed in Table 2.

When possible, the data layers in Table 2 used to derive basin parameter estimates are identical to those used in the development of the original regression equations. However, there are some parameters for which the *AvThreshR* database does not contain the exact data source used by the regression equation developers or the exact source used by the regression equation developers is not known. For example, it is often stated in Jennings et al. (1994) that parameters like surface water storage and percent forest cover can be derived from topographic maps, and the user should use the highest resolution maps available. In these situations, Endres (1999) collected data sources that were readily available in digital format.

DEM data and DEM derivatives are used to define basin boundaries and compute topographic parameters. DEM data with national coverage and a 15 arcsec spatial resolution were obtained from NOHRSC. This data set was created at NOHRSC by resampling 3 arcsec DEMs distributed by the USGS. For use in *AvThreshR*, DEM data sets (one for each RFC) provided by NOHRSC in geographic coordinates were projected into an Albers equal-area conic map projection and resampled to a 400-m resolution. Flow direction grids for each RFC were derived in projected space using information in the 400-m DEM and a modified version of the Environmental Protection Agency's River Reach File 1 (RF1) developed at NOHRSC ("Database" 2000). Flow accumulation, flow length, and slope grids were derived using Environmental Systems Research Institute (ESRI) grid functions (ESRI 1998a).

Some 30-m DEM data are also used to derive results presented





**Fig. 3.** Chart of: (a) stream catchments and longest flowpath lines traced for individual headwaters; (b) longest flowpath identified for nonheadwater basin 9. An example portion of a basin attribute table is also shown.

in this paper. This is possible because the *AvThreshR* algorithms are independent of DEM resolution. The required 30-m DEMs and DEM derivatives (e.g., flow direction) were provided by the National Severe Storms Laboratory (NSSL) ("National" 2000).

## Methods

### Threshold Runoff Calculations

Deriving a threshold runoff estimate for a basin can be divided into the following five basic steps:

1. Define the basin boundary.
2. Calculate physiographic and climatic parameters.
3. Compute  $Q_2$  using the appropriate regression equation for the location of interest (cms or cfs) (all required parameters are computed automatically by *AvThreshR*).
4. Compute the unit hydrograph peak flow ( $\text{cms} \cdot \text{mm}^{-1}$  or  $\text{cfs} \cdot \text{in.}^{-1}$ ) (topographic parameters are computed automatically, but a user must provide appropriate unit hydrograph coefficients).
5. Compute threshold runoff (mm or in.) by dividing the result from step 3 by the result from step 4 [Eq. (1)].

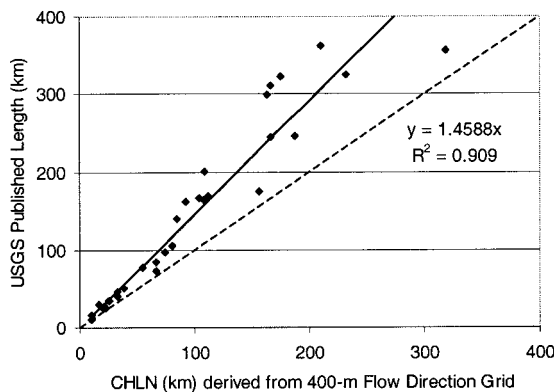
*AvThreshR* provides an efficient computational tool for estimating threshold runoff values for thousands of small basins within an RFC service area. The first step, defining basin boundaries, is straightforward, given the preprocessed DEM derivatives. The methods described by Jenson and Domingue (1988) are used, via ArcView functions, to define a network of synthetic streams from the flow direction grid (a grid in which an integer value in each cell indicates flow direction to one of its eight neighbors). This is done by first using the flow accumulation grid (in which each cell contains the number of upstream cells or drainage area) to classify cells as either stream cells or nonstream cells, depending on their upstream drainage area. A "stream catchment" or

**Table 3.** Basic Parameter Fields Added to Subbasin Attribute Table

Field name	Description
Lelvmean	Local mean elevation (m)
Lelvmax	Local maximum elevation (m)
Lelvmin	Local minimum elevation (m)
Lslpmean	Local mean slope (fraction)
Lflmean	Local mean flowlength (m)
Lflmax	Local maximum flowlength (m)
Lflmin	Local minimum flowlength (m)
Famax	Maximum flow accumulation
Lemaxlfp	Local maximum elevation along longest flow path (m)
Arm	Cumulative drainage area ( $\text{mi}^2$ )
Chsl0	Slope of longest flowpath (100%) (ft/mi), computed as difference in maximum and minimum elevations divided by longest flow path length
Chsl	"Channel slope," computed as difference in elevations at 85 and 10% along longest flowpath divided by distance between these points (ft/mi)
Chln	"Channel length" (mi), computed as longest flowpath from basin divide to outlet point
Chcn	"Channel centroid" or length to point on longest flowpath opposite basin centroid (mi)
Altind	Altitude index, computed as average of elevations at 10 and 85% along longest flowpath (thousands of ft)
Bshp	Basin shape = $\text{Chln}^2/\text{Arm}$
Centlat	Latitude of basin centroid
Centlon	Longitude of basin centroid
Shdishead	Code indicating whether subbasin is headwater (1 = headwater; 0 = nonheadwater)
Shdisout	Code indicating whether subbasin is outlet (has no subbasins downstream)
Stateabbr	Two-letter state code indicating which state contains subbasin centroid
Regions	List of regions intersected by cumulative subbasin shape
Reg_fract	Fraction of subbasin (within state of interest) that falls in each region

basin can be defined for each synthetic stream segment as shown in Fig. 3(a). *AvThreshR* also provides the options to (1) retain only headwater basins in the network for analysis; (2) specify an upper limit on basin size in addition to the lower limit defined by the stream threshold; and (3) allow the user to manually select basin outlet points. Similar basin delineation tools are widely used in the hydrologic modeling community (Maidment and Djokic 2000).

As stated earlier, some basin parameters are computed for all locations, while state-specific parameters (parameters only required for regressions in certain states) are only computed on an as-needed basis. A summary of the "basic" parameters computed for all subbasins is given in Table 3. Delineated basin boundaries are stored in polygon shapefiles (ESRI 1998b). Computed basin parameter values are stored in the attribute table associated with the shapefile (basin attribute table). The basin attribute table contains one record that corresponds to each watershed polygon, as illustrated in Fig. 3.



**Fig. 4.** *AvThreshR* derived channel length using 400-m DEM versus USGS published channel length for basins in Nebraska (Beckman 1976). The regression equation shown is used to adjust channel lengths when the 400-m DEM is used to compute lengths.

For nonheadwater or downstream polygons [5, 6, 8, and 9 in Fig. 3(a)], “local” parameters differ from cumulative parameters. For efficiency, the application computes local parameters for all basins first, followed by cumulative computations for nonheadwaters. It is the cumulative parameters that are required to compute results at a basin outlet point [e.g., cumulative area, average rainfall over the cumulative area, or cumulative flow length to the farthest upstream point as shown in Fig. 3(b) for subarea 9]. Calculating cumulative parameters requires information about basin connectivity, which is computed and stored by *AvThreshR*.

Because the USGS equations are applicable on a state-by-state, region-by-region basis, a basin of interest can intersect with more than one hydrologic region. When a basin intersects more than one hydrologic region within a given state, *AvThreshR* computes the fractional area of the basin in each region and area-weighted flow calculations are used. If a basin intersects more than one state, the location of the basin centroid is used to determine which state’s regression equations will be used.

Many of the studies summarized by Jennings et al. (1994) were done prior to the widespread availability of GIS software and data; therefore, basin parameters were measured using manual methods applied to paper map sources. *AvThreshR* attempts to derive parameters from digital sources using methods that produce similar results to the manual methods used by the USGS.

For example, channel length (*CHLN*) or longest flow path length is measured by USGS investigators as “the length of the main channel between the gaging station and the basin divide measured along the channel which drains the largest area” (Beckman 1976). Choquette (1987), Clement (1987), and Neely (1987) give the same definition for channel length. *AvThreshR* approximates channel length in the following manner. The flow length grid in the *AvThreshR* database stores the estimated distance from each DEM cell to the edge of the grid following the paths defined by the flow direction grid. The length of the path through a given cell is  $s$ ,  $\sqrt{2}s$ , or  $(s + s\sqrt{2})/2$ , where  $s$  is a cell side length. Computing the channel length for a basin is simply a matter of subtracting the minimum flow length in that basin from the maximum flow length. However, the accumulation of lengths along the cell-to-cell flow path will not exactly match the length of the natural channel or the blue line representation of the channel on a topographic map, and different channel lengths will be derived when different resolution DEMs are used. A plot of DEM derived

channel lengths versus USGS reported channel lengths for basins in Nebraska is shown in Fig. 4. Length estimates computed from the 400-m grid tend to be lower than those derived from paper maps by Beckman (1976). The linear regression equation shown in Fig. 4 is used in *AvThreshR* to adjust lengths ( $x$ ) computed from the DEM to yield more suitable values for the regression equations. This relationship should probably be varied for different types of terrain.

To account for differences among the required regression parameters for each RFC, a unique parameter code table (*parcode.dbf*) has been created for each RFC that tells the program how to compute specific parameters. The parameter code table includes information that indicates whether a basin average calculation is required or whether some other type of spatial calculation is required, where to find the input data in the database, what type of spatial data is used for input (grid or polygon), and how to write the computed parameter information to the basin attribute table. There is also a field in the parameter code table that contains unit conversion factors to be applied before substituting the results into a regression equation. Another table, the regression equation table (*regequat.dbf*), contains parameter names and coefficients defining the regression equations for each state, region, and return period.

$Q_2$  calculations are accomplished in two main steps. In the first step, information in the parameter code table is used along with the GIS data layers described in Tables 1 and 2 to derive basin parameter information. The results are stored in the basin attribute table (like that shown in Fig. 3). In the second step, parameter values from the basin attribute table and equation coefficients stored in the regression equation table are used to compute  $Q_2$ . More detailed specifications for the parameter code table and regression equation table are provided in the online documentation (“Threshold” 2001).

With a few exceptions, most of the state-specific parameters derived from the data layers listed in Table 2 are computed simply as an average of grid cells within a subbasin or as a weighted average of polygon pieces intersecting a subbasin. For example, mean annual precipitation is typically stored in a grid format, so the subbasin average is simply an average of grid cells. Surface water bodies are typically stored as polygons, so the percent storage in a subbasin is simply a weighted average of water storage polygons, where the storage areas are assigned a weight of 100 and the nonstorage areas are assigned a weight of 0. An example of a more complex parameter calculation supported by the code is the percent of the main channel that passes through swamp, a parameter required in the state of Michigan.

The Snyder unit hydrograph method (Snyder 1938) is used to estimate 1, 3, and 6 h unit hydrograph peaks for each subbasin. Carpenter et al. (1999) used both the Snyder unit hydrograph method (Snyder 1938) and the geomorphologic unit hydrograph (GUH) (Rodriguez-Iturbe and Valdes 1979) in their analysis. In the development of *AvThreshR*, it was assumed that estimates of Snyder coefficients are more readily available than relationships required to estimate GUH parameters; hence, the Snyder unit hydrograph is the default method programmed into *AvThreshR* and the Snyder method is used to compute results presented in this paper.

The Snyder method allows a synthetic unit hydrograph to be generated at an ungauged location using regional coefficients determined from gauged locations in the same region. In the Snyder method

$$Q_{pR} = \frac{C_1 C_p A}{0.9545 t_p + 0.25 t_R} \quad (3)$$

is used to estimate the unit hydrograph peak flow [ $\text{m}^3 \cdot \text{s}^{-1} \cdot \text{mm}^{-1}$  ( $\text{ft}^3 \cdot \text{s}^{-1} \cdot \text{in}^{-1}$ )] where  $C_1$  = a constant (0.278 for metric units and 645 for English units);  $C_p$  = empirical coefficient;  $t_p$  = lag time (h) (time difference between center of the rainfall duration and unit hydrograph peak); and  $t_R$  = rainfall duration. One of two equations is used in *AvThreshR* to estimate lag time

$$t_p = C_2 C_{t1} (L * L_c)^{0.3} \text{ [Synder (1938)]} \quad (4)$$

$$t_p = C_3 C_{t2} \left[ \frac{(L * L_c)}{(S)^{0.5}} \right]^{0.38} \text{ [“Flood” (1959)]} \quad (5)$$

where  $L$  = longest flow path length in the basin;  $L_c$  = length from the outlet to the point on the channel opposite the centroid; and  $S$  = dimensionless slope.  $C_{t1}$  and  $C_{t2}$  are the Snyder coefficients, often accepted as constants for a given region.  $C_2$  and  $C_3$  are unit conversion constants [metric (km):  $C_2 = 0.752$ ,  $C_3 = 0.697$ ; English (mi):  $C_2 = C_3 = 1$ ]. A user chooses between Eqs. (4) and (5), depending on the coefficient information that is available for the area of interest.

### Uncertainty Analysis

Using basin data from Oklahoma, an uncertainty analysis is done to approximate lower and upper error bounds within which there is a 68% probability that threshold runoff ( $R$ ) estimates for a given site will fall (equivalent to one standard deviation error bars for a normal distribution). The methods used to do this are described here, and results (in the form of error bounds) are given in the next section. After setting  $Q_f$  equal to  $Q_2$  and taking the logarithm of Eq. (1),

$$\ln R = \ln Q_2 - \ln Q_p \quad (6)$$

a simple analytical expression can be written to relate the variance of  $\ln R$  to the variances and covariance of  $\ln Q_2$  and  $\ln Q_p$

$$\text{Var}[\ln R] = \text{Var}[\ln Q_2] + \text{Var}[\ln Q_p] - 2 \text{Cov}[\ln Q_2, \ln Q_p] \quad (7)$$

Estimates of  $\text{var}[\ln Q_2]$  can be obtained from USGS reports describing regression equation development. The following equation is the  $Q_2$  equation for rural, unregulated basins in Oklahoma given in Jennings et al. (1994) [originally derived by Tortorelli and Bergman (1985)]:

$$Q_2 = a A^{0.59} P^{1.84} \quad (8)$$

and

$$\ln Q_2 = \ln(a) + 0.59 \ln A + 1.84 \ln P \quad (9)$$

is the log-transform of this equation. In these equations,  $A$  = drainage area [ $\text{km}^2$  ( $\text{mi}^2$ )];  $P$  = mean annual precipitation [ $\text{cm}$  ( $\text{in.}$ )]; and  $a$  = a constant (0.00107 for metric units and 0.368 for English units). Tortorelli and Bergman report the percent standard error ( $SE$ ) of the  $Q_2$  model as 60%. The percent  $SE$  is equivalent to the coefficient of variation ( $CV$ ) of the  $Q_2$  estimates times 100, or

$$SE' = CV = 100 * [(\text{Var}[Q_2])^{0.5} / E[Q_2]] \quad (10)$$

To apply Eq. (7), the variance of the estimate in log space,  $\text{var}[\ln Q_2]$ , is back-calculated from the  $CV$  using the following equation (Vanmarcke 1983):

$$\text{var}[\ln Q_2] = \ln[1 + (CV/100)^2] \quad (11)$$

For example, substituting  $CV = 60$  into Eq. (11) gives  $\text{var}[\ln Q_2] = 0.3075$ . In our uncertainty analysis, only the model error associated with applying the regional regression equations is considered, not the error associated with estimating  $Q_2$  values at individual sites from historical flood peak records.

Unit hydrograph peak flow ( $Q_p$ ) and time to peak ( $t_p$ ) data developed at the Tulsa Corps of Engineers Office (COE) and obtained from the Arkansas–Red Basin River Forecast Center (ABRFC) allow estimation of the  $\text{var}[\ln Q_p]$  and  $\text{Cov}[\ln Q_2, \ln Q_p]$  terms in Eq. (7). Site-specific estimates of  $Q_p$ ,  $t_p$ ,  $C_p$ , and  $C_{t2}$  for 54 basins were derived through hydrologic modeling at COE Tulsa. Length ( $L$ ), length to centroid ( $L_c$ ), and slope ( $S$ ) data were also provided for these basins. The mean of the  $C_p$  values for these basins, 0.75, is used as our regional estimate for  $C_p$ . The regional  $C_{t2}$  estimate, 0.279, is approximated as the slope of the line of best fit from a plot of time to peak ( $t_p$ ) against

$$C_3 [(L * L_c) / S^{0.5}]^{0.38} \text{ [see Eq. (5)].}$$

Unit hydrograph peak estimates are computed for all basins using Eq. (3) and the regional values of  $C_p$  and  $C_{t2}$ . The required model variance,  $\text{var}[\ln Q_p]$ , is computed as the sample variance of the differences ( $\ln \hat{Q}_p - \ln Q_p$ ) for all basins used to develop the model, where  $\hat{Q}_p$  is the unit hydrograph peak estimate derived using Eq. (3) and  $Q_p$  is the “true” value derived from hydrologic modeling. As with the specification of error for  $Q_2$  estimates, we only consider errors in  $Q_p$  estimation due to regional regression, not errors in deriving basin-specific  $Q_p$  values from hydrologic modeling. The derived value for  $\text{var}[\ln Q_p]$  is 0.1047. This corresponds to a  $CV$  for  $Q_p$  of 33%.

For 31 basins in the COE Tulsa data set,  $Q_2$  values originally derived from historical peak flood analysis were extracted from “WATSTORE” (2001) and  $\hat{Q}_2$  values were estimated using Eq. (8). With the values of both ( $\ln \hat{Q}_p - \ln Q_p$ ) and ( $\ln \hat{Q}_2 - \ln Q_2$ ) available for these 31 basins, the sample covariance of model errors,  $\text{Cov}[\ln Q_2, \ln Q_p]$ , is computed to be 0.0445. Substituting our derived values of  $\text{var}[\ln Q_2]$ ,  $\text{var}[\ln Q_p]$ , and  $\text{Cov}[\ln Q_2, \ln Q_p]$  into Eq. (7) yields  $\text{var}[\ln R] = 0.3232$ .

For given estimates of  $Q_2$  and  $Q_p$ , known uncertainty parameter ( $\text{var}[\ln R]$ ,  $\text{var}[\ln Q_2]$ , and  $\text{var}[\ln Q_p]$ ) values are used to approximate the probability bounds  $R_{16}$  [ $\text{Prob}(R < R_{16}) = 15.85\%$ ] and  $R_{84}$  [ $\text{Prob}(R < R_{84}) = 84.15\%$ ].  $Q_2$ ,  $Q_p$ , and  $R$  estimates for a given site all follow a lognormal distribution. Error bounds for the lognormal variable,  $R$ , are estimated using the Microsoft Excel function NORMSINV, which returns the inverse of the lognormal probability distribution given the desired probability level (0.1585 or 0.8415, in this case),  $E[\ln R]$ , and  $(\text{var}[\ln R])^{0.5}$ .  $E[\ln R]$  is calculated by first getting  $E[\ln Q_2]$  and  $E[\ln Q_p]$  by solving (Vanmarcke 1983)

$$E[X] = \exp\{E[\ln X] + 0.5 \text{var}[\ln X]\} \quad (12)$$

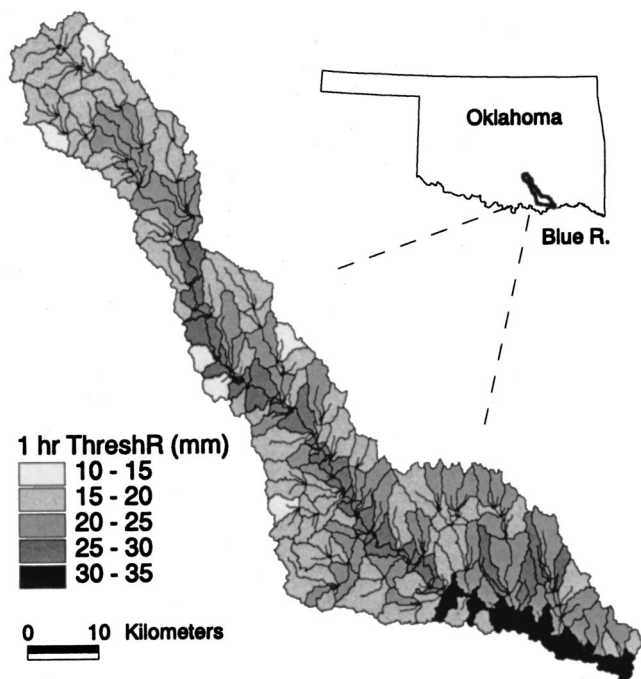
for  $E[\ln X]$  ( $X$  is a random variable with a lognormal distribution), and then using

$$E[\ln R] = E[\ln Q_2] - E[\ln Q_p] \quad (13)$$

### Results and Discussion

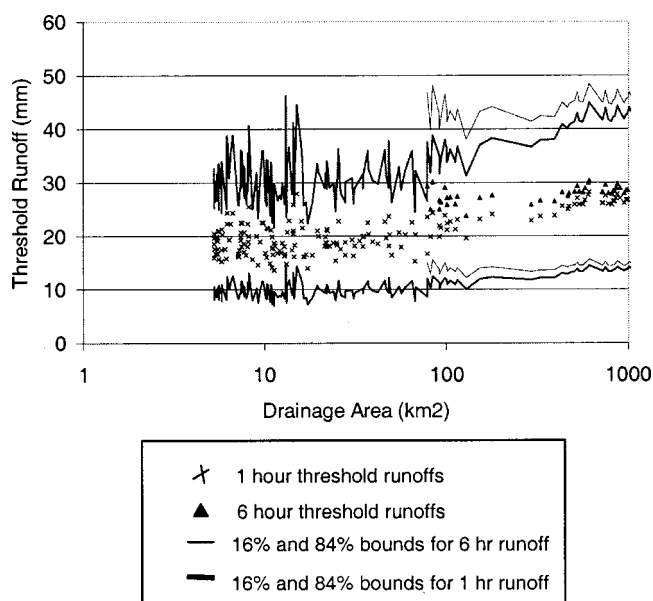
Threshold runoff analysis covering large areas (on the order of one million square kilometers) was done using the 400-m *AvThreshR* DEM derivatives, and analysis covering smaller areas (on the order of 1,000  $\text{km}^2$ ) was done using 30-m DEM derivatives. The reasons for analyzing 30-m data in addition to the



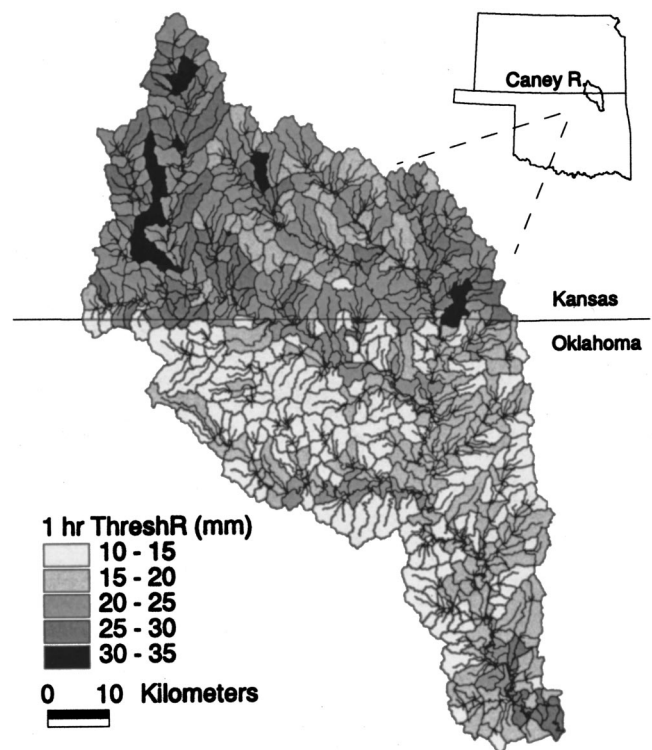


**Fig. 5.** One-hour threshold runoff in Blue River, Okla. Polygons are shaded based on computed values at the basin outlets. Values at outlets for nonheadwaters are based on cumulative upstream area.

400-m data that are part of the *AvThreshR* database are (1) to study threshold runoff values over a range of basin sizes (including small basins that cannot be accurately delineated using the 400-m data); and (2) to try to determine if using the 30-m data will help produce improvements to threshold runoff estimates that offset the additional computational and storage requirements associated with using 30-m data. A 30-m grid is huge in size relative to a 400-m grid covering the same area (approximately 178 times as much data).



**Fig. 6.** Drainage area versus threshold runoff in Blue River, Okla.



**Fig. 7.** One-hour threshold runoff in Caney River Basin

Using regional values of  $C_p$  (0.75) and  $C_{t2}$  (0.279) to estimate  $Q_p$  and using  $Q_2$  as  $Q_f$ , 1 and 6 h threshold runoff values were computed for subbasins in the Blue River Basin in Oklahoma (USGS HUC number 11140102) and the Caney River Basin in Oklahoma and Arkansas (USGS HUC number 11070106). A map of 1 h threshold runoff values for the Blue River Basin is shown in Fig. 5. The subbasins displayed in Fig. 5 were defined by first creating a synthetic stream network using a 5 km<sup>2</sup> threshold and then defining a subbasin for each synthetic stream reach. The 30-m DEM derivatives obtained from the NSSL were used for this analysis. Subbasins are shaded according to the threshold runoff values computed at their respective subbasin outlets based on cumulative drainage areas. The somewhat higher threshold runoff values and darker colors correspond to subbasins along the main channel with very large cumulative drainage areas. A better sense of the variation of threshold runoff with drainage areas can be seen in a plot of drainage area versus 1 and 6 h runoffs (Fig. 6). Fig. 6 shows that the 1 and 6 h runoffs are fairly insensitive to drainage area, with only a mild increasing trend for the range of basin sizes considered. The 6 h runoff values are only plotted for basins larger than 78 km<sup>2</sup> (30 mi<sup>2</sup>) because below this basin size, the time between the start of rainfall and unit hydrograph peak [denominator of Eq. (3)] can be less than 6 h, at which point the 6 h unit hydrograph becomes meaningless.

Physical reasoning can lead to the conclusion that both  $Q_f$  and  $Q_p$  should increase proportionally to drainage area raised to a power between 0 and 1; however, it is not obvious from this reasoning whether the threshold runoff quotient should increase or decrease with drainage area. Therefore, the runoff variations with drainage area shown in Fig. 6 are simply a function of the empirical relationships that we have assumed to be valid in our analysis. Leopold et al. (1992) assert that the empirical observations showing that channel capacity ( $Q_b$ ) increases proportionally to area raised to some power less than one are sensible because



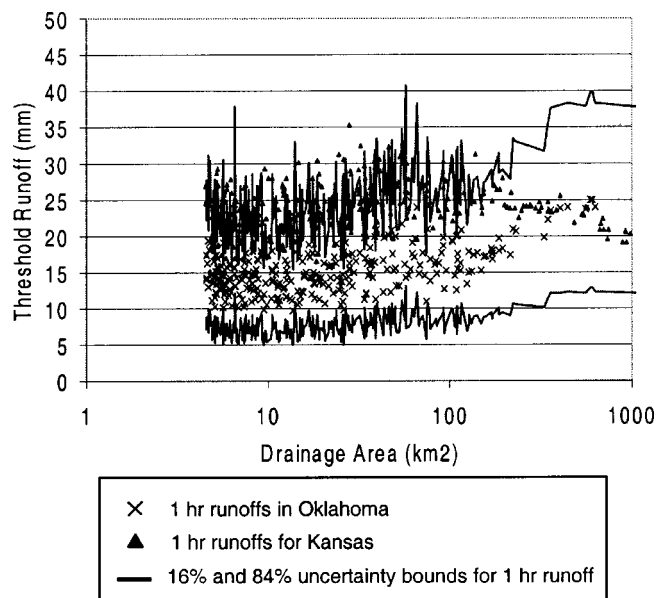


Fig. 8. Drainage area versus threshold runoff for Caney River Basin

the intensity of rainstorms decreases if rainfall is averaged over larger drainage areas and flood waves have more time to diffuse in larger basins. It is also reasonable to assume that the unit hydrograph peak ( $Q_p$ ) will continue to increase with drainage area, but that the unit graph peak per unit area ( $q_p$ ) should decrease with basin size due to the effects of diffusion/dispersion.

The 16 and 84% uncertainty bounds for 1 and 6 h runoffs are drawn in Fig. 6. It is noteworthy that the error bounds on the 1 h runoff values easily encompass the deterministic estimates for 6 h runoff. The spread of these error bounds is quite large, and the spread of the error bounds increases with increasing threshold runoff values. In fact, the difference between the 16 and 84%

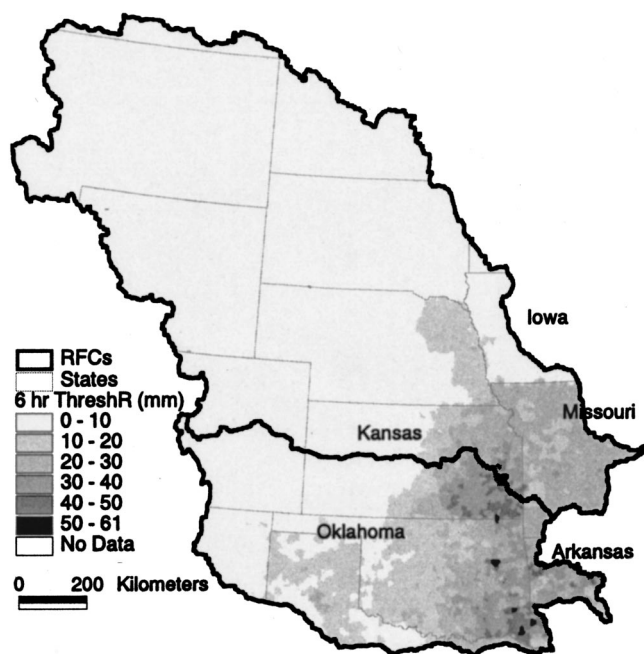


Fig. 9. Initial 6-h threshold runoff estimates for MBRFC and ABRFC

bounds for a single point (e.g.  $R_{1h}=23$  mm;  $R_{16}=11.9$  mm;  $R_{84}=37.1$  mm) can be greater than the total range of 1 h deterministic estimates shown in Fig. 6 for all basins, which is 21.5 mm.

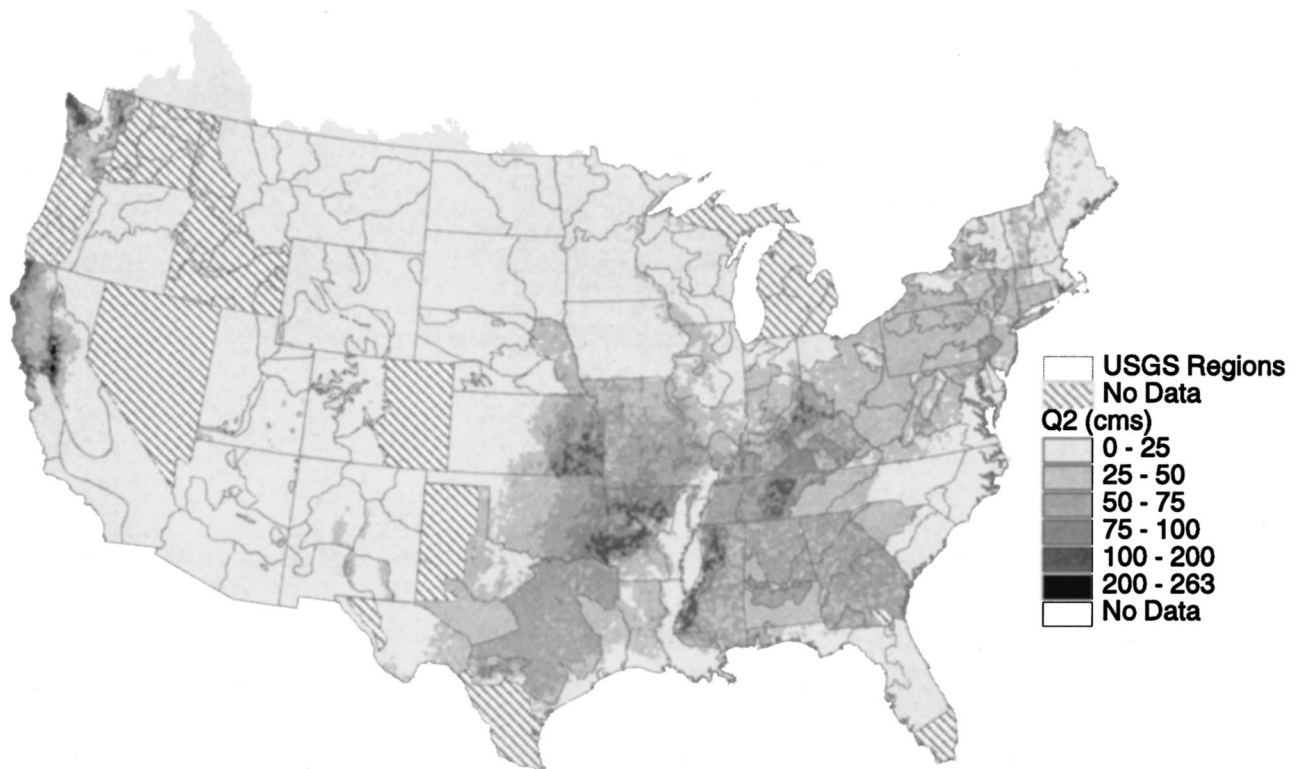
It turns out that another important factor to consider when using USGS  $Q_2$  equations is the spatial variability of estimates from state to state. A map of 1 h threshold runoff values for the Caney River Basin is shown in Fig. 7. An abrupt shift in the magnitude of the threshold runoff values is seen across the border of Kansas and Oklahoma. This is caused by the fact that Kansas regression equations predict higher  $Q_2$  values than the Oklahoma equations for basins with similar characteristics. A plot of drainage area versus 1 h runoffs for the Caney River Basin is shown in Fig. 8. Basins in Oklahoma in the size range of 10–100 km<sup>2</sup> have an average 1 h runoff of 15.1 mm, while the average for basins of this size just across the border in Kansas is 24.6 mm. In Fig. 8, 16 and 84% error bounds are provided for the Oklahoma basins only. For basins smaller than about 150 km<sup>2</sup>, 1 h threshold runoffs for Kansas can lie above the 84% error bound for Oklahoma data. This indicates that state-to-state inconsistencies are at least as significant as, and perhaps more significant than, uncertainty due to the use of regional  $Q_2$  regressions and Snyder coefficients.

To give a sense of how threshold runoff varies over large regions, Fig. 9 shows initial 6 h threshold runoffs computed for the ABRFC service area and the Missouri Basin River Forecast Center (MBRFC) service area. The total service areas for these RFCs cover 1,895,000 km<sup>2</sup> (732,000 mi<sup>2</sup>). To derive the results shown in Fig. 9, 7,105 headwater basins ranging in size from 78 to 130 km<sup>2</sup> (30–50 mi<sup>2</sup>) were delineated and threshold runoff values were computed for each of these basins. To create the 4 by 4 km gridded surface shown here, gaps not covered by these headwaters were filled in using inverse distance weighting interpolation. Headwater basins of this size are representative of flash-flood scale basins.

To facilitate the creation of the threshold runoff surface shown in Fig. 9, we made a liberal assumption that the regional Snyder coefficients are constant over this entire area. Initial attempts to derive spatially varying grids of  $C_p$  and  $C_{t2}$  values by interpolating values from known locations did not yield sensible spatial patterns. A more satisfactory, yet more complex, solution to derive spatially distributed Snyder coefficients would be to relate  $C_p$  and  $C_{t2}$  to watershed characteristics, as done by Miller et al. (1983) and the Hydrologic Engineering Center (1989).

Based on local knowledge and past experience, hydrologists at ABRFC and MBRFC have indicated that several patches of the threshold runoff values in Fig. 9 are too high (Robert Cox, personal communication, 2000; Billy Olsen, personal communication, 2000). The highest 6 h threshold runoff values that are currently used in the ABRFC and MBRFC areas are between 38 and 51 mm (1.5–2.0 in.); however, there are several patches in Fig. 9 that exceed these values. The unexpectedly high patches of threshold runoff occur primarily in southeastern Kansas, southeastern Oklahoma, and western Arkansas. In addition to the high threshold runoff values, there are some noticeable shifts in threshold runoff magnitudes at state borders (Kansas-Oklahoma, Kansas-Missouri, and Missouri-Iowa borders, in particular). These shifts at state borders are due to the use of  $Q_2$  regression equations that differ for neighboring states, as discussed in connection with Figs. 7 and 8.

There are a number of reasons why abrupt shifts in  $Q_2$  magnitudes may occur across state boundaries. The investigators who derived USGS flood frequency equations for different states did their studies independently, at different times, and using different



**Fig. 10.**  $Q_2$  estimates in conterminous U.S. derived for basins ranging in size from 52 to 91 km<sup>2</sup> (20–35 mi<sup>2</sup>)

periods of streamflow data. For the rural equations, the dates of the original publications range from 1973 to 1993. The statistical techniques for deriving flood frequency equations have also evolved over time (Jennings et al. 1994).

Independent flood frequency studies also resulted in different methods for dealing with spatial variability. As shown in Fig. 2, some states are divided into multiple regions where different equations apply. The partitioning of states into separate regions is one way to account for spatial variability without explicitly including additional parameters besides drainage area in the regression equations. For example, Tennessee is divided into four hydrologic regions, but the regression equations for all four regions are only a function of drainage area. Oklahoma is treated as one region, but mean annual precipitation appears in the regression equations to account for climatic variability. In other states, various combinations of regional definitions, physiographic, and climatic parameters are used to account for spatial variability.

In an effort to determine the degree to which  $Q_2$  magnitude shifts at regional and state boundaries may produce irregularities in computed threshold runoff values, a national map of  $Q_2$  for basins ranging in size from 52 to 91 km<sup>2</sup> (20–35 mi<sup>2</sup>) was created (Fig. 10). Like in Fig. 9, the  $Q_2$  values computed for these basins were interpolated to a regular grid for display. In Fig. 10, regions where the  $Q_2$  values could not be calculated are denoted by hash marks (the same hashed areas highlighted in Fig. 2 with the addition of one area in western Oregon where we have a corrupted data layer). Relatively high  $Q_2$  estimates are apparent in southeastern Kansas, southeastern Oklahoma, parts of Arkansas, parts of Mississippi, middle Tennessee, parts of Kentucky, and California near the San Francisco Bay. Without a more detailed study, it is not possible to identify the causes of these higher  $Q_2$  values and determine whether these anomalies reflect reality or whether they are a by-product of differences in regression equation devel-

opment procedures in different states. The USGS is in the process of updating regression equations (Ries, personal communication, 2000), so it will be interesting to see if regional discrepancies are smoothed with these new equations. It is also likely that the use of updated regression equations can shrink the error bounds plotted in Figs. 6 and 8 because the standard error of 60% reported for the Oklahoma equations is relatively high and many states report lower standard errors, down to about 30%.

## Conclusions

A variety of methods have been used in the past to derive the countywide threshold runoff values currently used to produce flash flood guidance estimates. Therefore, a more consistent, objective, and automated procedure is sought to estimate threshold runoff values over RFC service areas. Development of GRASS-based threshold runoff software (*threshR*) in 1993 (Carpenter and Georgakakos 1993) demonstrated the potential for objectively and automatically deriving threshold runoff estimates for basins of various sizes. In 1999, collection, processing, and development of supporting data sets were initiated in an effort to further automate flooding flow estimation (a key input to threshold runoff calculations) and reduce the computational burden of GIS-based procedures at RFCs. The resulting ArcView-based application (*AvThreshR*) can now support the automatic estimation of 2, 5, and 10-year floods, etc., based on U.S. Geological Survey regression equations for rural areas.

The basin delineation accuracy of the 400-m terrain database delivered with *AvThreshR* degrades for smaller and smaller basins [a rule of thumb is not to delineate basins smaller than 78 km<sup>2</sup> (30 mi<sup>2</sup>)]. Although it was initially assumed that making threshold runoff calculations on basins smaller than 78 km<sup>2</sup> might be im-

portant, results from this study indicate that the increase in threshold runoff values with drainage areas is relatively small compared to the uncertainty inherent in using the combination of USGS  $Q_2$  equations to estimate  $Q_f$  and the Snyder method to define  $Q_p$ . Therefore, until research produces more accurate regional equations or better methods to reduce the uncertainty bounds on threshold runoffs for basins of all sizes, attempts to compute threshold runoffs on very small basins may not provide much benefit. It follows that the use of a relatively coarse 400-m DEM in initial screening has proven to be sensible.

A key assumption in our analysis is that the  $Q_2$  estimates produced by the USGS regression equations are a useful surrogate for flooding flow. Based on the experience of RFC hydrologists, initial  $Q_2$  estimates in some areas are too high to use as practical flooding flow estimates. This may be due to the uncertainty in applying regression relationships to estimate  $Q_2$  and/or the assumption that  $Q_2$  is a good surrogate for flooding flow in a diverse set of basins. An additional difficulty in using the current  $Q_2$  regression equations is the presence of abrupt shifts in  $Q_2$  estimates that occur across some state and regional boundaries. It is possible that updated regression equations would alleviate some of these issues. Given the *AvThreshR* capabilities to derive numerous basin parameters and implement numerous regression equations, testing new equations would be relatively easy.

There are certain limitations inherent in using automated GIS parameter estimation procedures if the resulting parameter estimates do not match manually derived parameters used to develop the empirical regression equations being used. For example, DEM-based length and slope measurements can be systematically different from manually estimated values. When possible, empirical corrections have been made to the DEM computed parameter values. It would be advantageous in the future development of regression equations to base parameter estimates on nationally available digital databases.

Carpenter et al. (1999) discussed both statistically based and physically based approximations for flooding flow. The focus of the present study has been on the statistically based approach because of the nearly complete national coverage of  $Q_2$  regression equations that are currently available. In light of difficulties in using the existing  $Q_2$  equations for some parts of the country, it seems that the use of a more physically based approximation for flooding flow deserves further consideration. Hopefully, knowledge gained from ongoing work in distributed modeling, terrain analysis, and floodplain mapping at the NWS will lead to improved methods for estimating flooding flows at ungaged locations.

It should be kept in mind that threshold runoff is only one component of flash flood guidance. Distributed modeling research to address space-time discrepancies that exist between calibrated rainfall-runoff models and flash flood scale runoff processes must parallel efforts to improve flooding flow estimates.

## Acknowledgments

Karen Endres, a graduate student at Michigan Technological University, and Alex Metcalf, an undergraduate at Juniata College, collected, processed, and edited data sets for this project. Feedback from Eugene Derner and Robert Cox of the Missouri Basin River Forecast Center, and Michael Pierce and Billy Olsen of the Arkansas-Red Basin River Forecast Center was helpful in evaluating the initial results.

## References

- "Basin delineation with a 400-m terrain dataset." (<http://hsp.nws.noaa.gov/oh/hrl/gis/delin.html>) (November 13, 2001).
- Beckman, E. W. (1976). "Magnitude and frequency of floods in Nebraska." *USGS Water-Resources Investigations Rep. 76-109*, U.S. Geological Survey, Menlo Park, Calif.
- Carpenter, T. M., and Georgakakos, K. P. (1993). "GIS-based procedures in support of flash flood guidance." *Iowa Institute of Hydraulic Research Rep. No. 366*, Iowa City, Iowa.
- Carpenter, T. M., Sperflage, J. A., Georgakakos, K. P., Sweeney, T., and Fread, D. L. (1999). "National threshold runoff estimation utilizing GIS in support of operational flash flood warning systems." *J. Hydrol.*, 224(1-2), 21-44.
- Choquette, A. F. (1987). "Regionalization of peak discharges for streams in Kentucky." *USGS Water-Resources Investigations Rep. 87-4209*, U.S. Geological Survey, Menlo Park, Calif.
- Clement, R. W. (1987). "Floods in Kansas and techniques for estimating their magnitude and frequency on unregulated streams." *USGS Water-Resources Investigations Rep. 87-4008*, U.S. Geological Survey, Menlo Park, Calif.
- "Database preparation (Av-ThreshR documentation)." (<http://hsp.nws.noaa.gov/oh/hrl/gis/data.html#section214>) (Dec. 15, 2000).
- Davis, R. S. (1998). "Detecting time duration of rainfall: A controlling factor of flash flood intensity." *Special Symposium on Hydrology, American Meteorological Society Phoenix*, 258-263.
- Endres, K. L. (1999). "Compilation of GIS data layers for flash flood forecasting." Master's thesis, Dept. of Civil and Environmental Engineering, Michigan Technological Univ., Houghton, Mich.
- Environmental Systems Research Institute (ESRI). (1998a). "ArcDoc version 7.2.1—Patch 1." Rep., Redlands, Calif.
- Environmental Systems Research Institute (ESRI). (1998b). "ESRI shapefile technical description." *ESRI White Paper Rep.*, Redlands, Calif.
- Finnerty, B. D., Smith, M. B., Seo, D. J., Koren, V., and Moglen, G. E., (1997). "Space-time scale sensitivity of the Sacramento model to radar-gage precipitation inputs." *J. Hydrol.*, 203, 21-38.
- "Flood hydrograph analysis and computations." (1959). *Engineering and design manual, EM 1110-2-1405*, U.S. Government Printing Office, Washington, D.C.
- GRASS 4.1 user's reference manual. (1983). U.S. Army Corps of Engineers Construction Engineering Research Laboratory, Champaign, Ill.
- Henderson, F. M. (1966). *Open channel flow*, Macmillan, New York.
- Hydrologic Engineering Center (HEC). (1989). "Real-time flood forecasting and reservoir regulation for the Allegheny River Basin." *Project Rep. No. 89-5*.
- Jennings, M. E., Thomas, W. O. Jr., and Riggs, H. C. (1994). "Nation-wide summary of U.S. Geological Survey regional regression equations for estimating magnitude and frequency of floods for ungaged sites, 1993." *USGS Water-Resources Investigations Rep. 94-4002*, U.S. Geological Survey, Menlo Park, Calif.
- Jenson, S. K., and Domingue, J. O. (1988). "Extracting topographic structure from digital elevation data for geographic information system analysis." *Photogramm. Eng. Remote Sens.*, 54(11), 1593-1600.
- Koren, V., Smith, M., Reed, S., and Zhang, Z. (2001). "Transition from lumped to distributed modeling systems." *Proc., AGU 2001 Spring Meeting, Presentation H32C-01*.
- Leopold, L., Wolman, M. G., and Miller, J. P. (1992). *Fluvial processes in geomorphology*, Dover, New York.
- Maidment, D. M., and Djokic, D. (2000). *Hydrologic and hydraulic modeling support with geographic information systems*, ESRI Press, Redlands, Calif.
- Miller, A. C., Kerr, S. N., and Spaeder, D. J. (1983). "Calibration of Snyder coefficients for Pennsylvania." *Water Resour. Bull.*, 625-630.
- "National basin delineation project." (<http://www.nssl.noaa.gov/teams/western/basins/>) (December 15, 2000).
- Neely, B. L. (1987). "Magnitude and frequency of floods in Arkansas." *USGS Water-Resources Investigations Rep. 86-4335*, U.S. Geological Survey, Menlo Park, Calif.

- Rodriguez-Iturbe, I., and Valdes, J. B., (1979). "The geomorphological structure of hydrologic response." *Water Resour. Res.*, 15(6), 1409–1420.
- Smith, M. B., Koren, V., Finnerty, B., and Johnson, D. (1999). "Distributed modeling: Phase 1 results." *NOAA Technical Rep. NWS 44*, National Oceanic and Atmospheric Administration, Spokane, Wash.
- Snyder, F. F. (1938). "Synthetic unit-graphs." *Trans., Am. Geophys. Union*, 19, 447–454.
- Sweeney, T. L. (1992). "Modernized areal flash flood guidance." *NOAA Technical Memorandum NWS Hydro 44*, National Oceanic and Atmospheric Administration, Spokane, Wash.
- "Threshold runoff." (<http://hsp.nws.noaa.gov/oh/hrl/gis/threshrhome.html>) (October 10, 2001).
- Tortorelli, R. L., and Bergman, D. L. (1985). "Techniques for estimating flood peak discharges for unregulated streams and streams regulated by small floodwater retarding structures in Oklahoma." *USGS Water-Resources Investigations Rep. 84-4358*, U.S. Geological Survey, Menlo Park, Calif.
- Vanmarcke, E. (1983). *Random fields: Analysis and synthesis*, MIT Press, Cambridge, Mass.
- "WATSTORE stream flow basin characteristics file." (<http://water.usgs.gov/GIS/metadata/usgswrd/sfbc.html>) (October 25, 2001).
- Wolman, M. G., and Leopold, L. B. (1957). "River flood plains: Some observations on their formation." *USGS Professional Paper 282-C*, U.S. Geological Survey, Menlo Park, Calif.